(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 3 125 152 A1

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
01.02.2017 Bulletin 2017/05

(51) Int Cl.:
G06K 9/00 *(2006.01)*

(21) Application number: **16179302.1**

(22) Date of filing: **13.07.2016**

(84) Designated Contracting States:
AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR
Designated Extension States:
BA ME
Designated Validation States:
MA MD

(30) Priority: **31.07.2015 CN 201510462792**

(71) Applicant: **Xiaomi Inc.**
Beijing 100085 (CN)

(72) Inventors:
- **TANG, Mingyong**
  Beijing 100085 (CN)
- **LIU, Huayijun**
  Beijing 100085 (CN)
- **CHEN, Tao**
  Beijing 100085 (CN)

(74) Representative: **Underwood, Nicolas Patrick et al**
Cabinet Beau de Loménie
158, rue de l'Université
75340 Paris Cedex 07 (FR)

(54) METHOD AND DEVICE FOR COLLECTING SOUNDS CORRESPONDING TO SURVEILLANCE IMAGES

(57) The present invention relates to a method and device for collecting sounds corresponding to surveillance images, which belongs to the field of computer technology. The method for collecting sounds corresponding to surveillance images includes: when surveillance images are collected by a camera of a monitoring device, detecting (101) whether there is a surveilled object making sounds in the surveillance images; if there is a surveilled object making sounds in the surveillance images, determining (102) a sound acquiring device meeting a preset condition which requires that a signal strength of the sounds collected by the sound acquiring device from the surveilled object is greater than a signal strength of the sounds collected by the monitoring device from the surveilled object; and controlling the sound acquiring device to collect the sounds from the surveilled object.invention

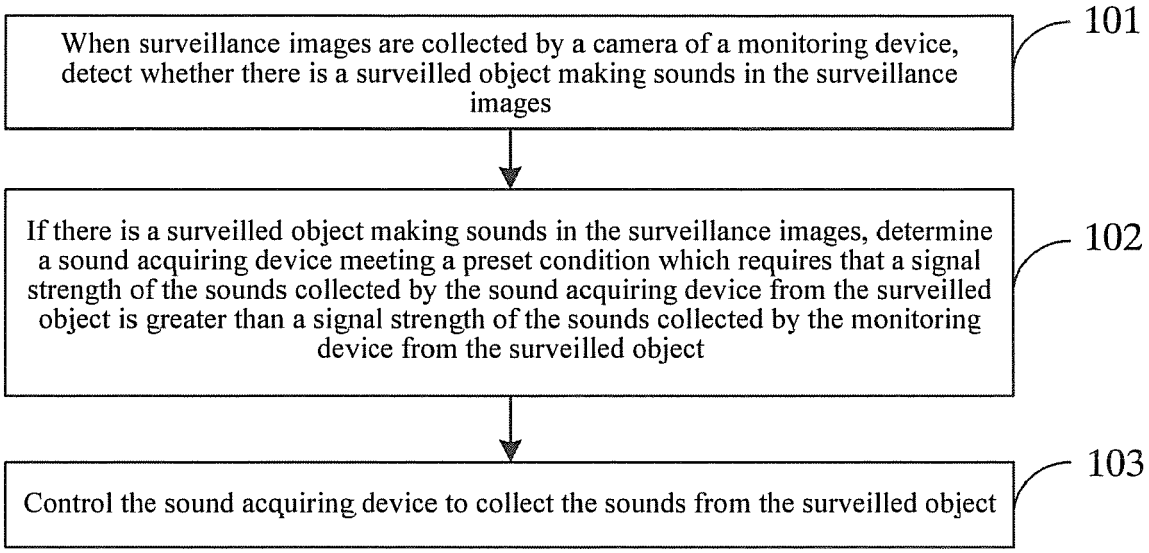

Fig. 1

EP 3 125 152 A1

Description

TECHNICAL FIELD

[0001] The present invention generally relates to the field of computer technologies, and more particularly, to a method and device for collecting sounds corresponding to surveillance images.

BACKGROUND

[0002] Cameras can truly monitor the procedure of various events, and the monitoring function of cameras is widely applied, for example, to traffic surveillance, to supermarket surveillance, or to community surveillance, and the like.

[0003] During surveillance using cameras on a monitoring device, if the surveilled object to be captured is relatively far, the monitoring device cannot clearly collect the sounds from the surveilled object, and thus the surveillance information is not complete.

SUMMARY

[0004] In order to address the problem related to a monitoring device not being able to clearly collect sounds from a surveilled object which is relatively far from the monitoring device, the present invention provides a method and a device for collecting sounds corresponding to surveillance images. The technical solutions are as follows.

[0005] According to a first aspect of embodiments of the present invention, there is provided a method for collecting sounds corresponding to surveillance images, including:

when surveillance images are collected by a camera of a monitoring device, detecting whether there is a surveilled object making sounds in the surveillance images;
if there is a surveilled object making sounds in the surveillance images, determining a sound acquiring device meeting a preset condition which requires that a signal strength of the sounds collected by the sound acquiring device from the surveilled object is greater than a signal strength of the sounds collected by the monitoring device from the surveilled object; and
controlling the sound acquiring device to collect the sounds from the surveilled object.

[0006] According to a particular embodiment, the detecting whether there is a surveilled object making sounds in the surveillance images, includes:

detecting whether there is an image changing element making sounds in the surveillance images, wherein the image changing element includes a surveilled object entering into the surveillance images or a surveilled object whose lips move; or
detecting whether an object of a designated type in the surveillance images makes sounds, wherein the object of the designated type includes at least one of a door, a window and a home appliance; or
according to a designated index reported by a detection device, detecting whether there is a surveilled object making sounds in the surveillance images, wherein the detection device is a device bound to the monitoring device for detecting the designated index, and the designated index includes one of on/off states, air quality parameters, temperature, parameters indicating changes of heat sources or physiological feature parameters.

[0007] According to a particular embodiment, the determining a sound acquiring device meeting a preset condition includes:

detecting whether sounds corresponding to the surveillance images collected by the monitoring device are smaller than a preset sound threshold; and
if the sounds corresponding to the surveillance images collected by the monitoring device are smaller than the preset sound threshold, determining the sound acquiring device meeting the preset condition.

[0008] According to a particular embodiment, the determining the sound acquiring device meeting the preset condition includes:

identifying a sound acquiring device closest to the surveilled object making sounds among all sound acquiring devices bound to the monitoring device, and determining the identified sound acquiring device as the sound acquiring device meeting the preset condition; or
controlling all sound acquiring devices bound to the monitoring device to collect the sounds from the surveilled object making sounds, and determining a sound acquiring device which collects a strongest sound signal as the sound acquiring device meeting the preset condition; or
determining at least one mobile device bound to the monitoring device as the sound acquiring device meeting the preset condition.

[0009] According to a particular embodiment, the identifying a sound acquiring device closest to the surveilled object making sounds among all sound acquiring devices bound to the monitoring device, includes:

by an indoor positioning technology, identifying the sound acquiring device closest to the surveilled object making sounds among all sound acquiring devices bound to the monitoring device; or
by an image recognition technology, determining a

location where the surveilled object making sounds is, and determining a sound acquiring device for monitoring the sounds from the surveilled object making sounds according to respective surveillance area ranges corresponding to the sound acquiring devices.

[0010] According to a particular embodiment, the controlling the sound acquiring device to collect the sounds from the surveilled object, includes:

when the sound acquiring device meeting the preset condition is a sound acquiring device closest to the surveilled object making sounds, controlling the sound acquiring device to collect sounds; or
when the sound acquiring device meeting the preset condition is a sound acquiring device which collects the strongest sound signal, controlling the sound acquiring device to collect sounds; or
when the sound acquiring device meeting the preset condition is a mobile device bound to the monitoring device, controlling the mobile device to move to a location of the surveilled object making sounds in the surveillance images, and controlling the mobile device to collect sounds at the location where the surveilled object is.

[0011] According to a second aspect of embodiments of the present invention, there is provided a device for collecting sounds corresponding to surveillance images, including:

a detection module configured to, when surveillance images are collected by a camera of a monitoring device, detect whether there is a surveilled object making sounds in the surveillance images;
a determination module configured to, if the detection module detects that there is a surveilled object making sounds in the surveillance images, determine a sound acquiring device meeting a preset condition which requires that a signal strength of the sounds collected by the sound acquiring device from the surveilled object is greater than a signal strength of the sounds collected by the monitoring device from the surveilled object; and
a control module configured to control the sound acquiring device determined by the determination module to collect the sounds from the surveilled object.

[0012] According to a particular embodiment, the detection module includes:

a first detection submodule configured to detect whether there is an image changing element making sounds in the surveillance images, wherein the image changing element incudes a surveilled object entering into the surveillance images or a surveilled object whose lips move; or
a second detection submodule configured to detect whether an object of a designated type in the surveillance images makes sounds, wherein the object of the designated type includes at least one of a door, a window and a home appliance; or
a third detection submodule configured to, according to a designated index reported by a detection device, detect whether there is a surveilled object making sounds in the surveillance images, wherein the detection device is a device bound to the monitoring device for detecting the designated index, and the designated index includes one of on/off states, air quality parameters, temperature, parameters indicating changes of heat sources or physiological feature parameters.

[0013] According to a particular embodiment, the determination module includes:

a fourth detection submodule configured to detect whether sounds corresponding to the surveillance images collected by the monitoring device are smaller than a preset sound threshold; and
a first determination submodule configured to, if the fourth detection submodule detects that the sounds corresponding to the surveillance images collected by the monitoring device are smaller than the preset sound threshold, determine the sound acquiring device meeting the preset condition.

[0014] According to a particular embodiment, the first determination submodule includes:

a first identifying submodule configured to identify a sound acquiring device closest to the surveilled object making sounds among all sound acquiring devices bound to the monitoring device, and determine the identified sound acquiring device as the sound acquiring device meeting the preset condition; or
a second identifying submodule configured to control all sound acquiring devices bound to the monitoring device to collect the sounds from the surveilled object making sounds, and determine a sound acquiring device which collects a strongest sound signal as the sound acquiring device meeting the preset condition; or
a second determination submodule configured to determine at least one mobile device bound to the monitoring device as the sound acquiring device meeting the preset condition.

[0015] According to a particular embodiment, the first identifying submodule includes:

a third identifying submodule configured to, by an indoor positioning technology, identify the sound acquiring device closest to the surveilled object making sounds among all sound acquiring devices bound to

the monitoring device; or

a fourth determination submodule configured to, by an image recognition technology, determine a location where the surveilled object making sounds is, and determine a sound acquiring device for monitoring the sounds from the surveilled object making sounds according to respective surveillance area ranges corresponding to the sound acquiring devices.

[0016] According to a particular embodiment, the control module includes:

a first control submodule configured to, when the sound acquiring device meeting the preset condition is a sound acquiring device closest to the surveilled object making sounds, control the sound acquiring device to collect sounds; or

a second control submodule configured to, when the sound acquiring device meeting the preset condition is a sound acquiring device which collects the strongest sound signal, control the sound acquiring device to collect sounds; or

a third control submodule configured to, when the sound acquiring device meeting the preset condition is a mobile device bound to the monitoring device, control the mobile device to move to a location of the surveilled object making sounds in the surveillance images, and control the mobile device to collect sounds at the location where the surveilled object is.

[0017] According to a third aspect of embodiments of the present invention, there is provided a device for collecting sounds corresponding to surveillance images, including:

a processor; and

a memory for storing instructions executable by the processor;

wherein the processor is configured to:

when surveillance images are collected by a camera of a monitoring device, detect whether there is a surveilled object making sounds in the surveillance images;

if there is a surveilled object making sounds in the surveillance images, determine a sound acquiring device meeting a preset condition which requires that a signal strength of the sounds collected by the sound acquiring device from the surveilled object is greater than a signal strength of the sounds collected by the monitoring device from the surveilled object; and

control the sound acquiring device to collect the sounds from the surveilled object.

[0018] In one particular embodiment, the steps of the method as defined above are determined by computer program instructions.

[0019] Consequently, according to a fourth aspect, the invention is also directed to a computer program which, when executed on a processor, performs a method as defined above.

[0020] This program can use any programming language and take the form of source code, object code or a code intermediate between source code and object code, such as a partially compiled form, or any other desirable form.

[0021] The invention is also directed to a computer-readable information medium containing instructions of a computer program as described above.

[0022] The information medium can be any entity or device capable of storing the program. For example, the medium can include storage means such as a ROM, for example a CD ROM or a microelectronic circuit ROM, or magnetic storage means, for example a diskette (floppy disk) or a hard disk.

[0023] Alternatively, the information medium can be an integrated circuit in which the program is incorporated, the circuit being adapted to execute the method in question or to be used in its execution.

[0024] The technical solutions provided by the embodiments of the present invention may have the following advantageous effects:

A surveilled object making sounds in the surveillance images is identified, and a sound acquiring device which collects sounds of a greater signal strength from the surveilled object than the signal strength of the sounds collected by the monitoring device from the surveilled object is controlled to collect the sounds from the surveilled object. Thus, the present invention can solve the problem related to a monitoring device not being able to clearly collect sounds from a surveilled object which is relatively far from the monitoring device. The monitoring device controls the sound acquiring device to collect the sounds from the surveilled object, and thus the surveillance information becomes more complete.

[0025] It is to be understood that both the foregoing general description and the following detailed description are exemplary only and are not restrictive of the present invention, as claimed.

BRIEF DESCRIPTION OF THE DRAWINGS

[0026] The accompanying drawings, which are incorporated in and constitute a part of this specification, illustrate embodiments consistent with the present invention and, together with the description, serve to explain the principles of the present invention.

Fig. 1 is a flowchart showing a method for collecting sounds corresponding to surveillance images according to an exemplary embodiment.

Fig. 2A is a flowchart showing detection of whether there is a surveilled object making sounds in the surveillance images according to an exemplary embodiment.

Fig. 2B is a flowchart showing a method for detecting whether there is a surveilled object making sounds in the surveillance images according to an exemplary embodiment.

Fig. 2C is a schematic diagram showing that a surveilled object enters into surveillance images according to an exemplary embodiment.

Fig. 2D is a schematic diagram showing that there is a surveilled object whose lips move in surveillance images according to an exemplary embodiment.

Fig. 2E is a schematic diagram showing an object of a designated type making sounds in the surveillance images according to an exemplary embodiment.

Fig. 2F is a schematic diagram showing detection of whether there is a surveilled object making sounds in the surveillance images according to a designated index reported by a detection device according to an exemplary embodiment.

Fig. 2G is a flowchart showing a method for determining a sound acquiring device meeting a preset condition according to an exemplary embodiment.

Fig. 2H is a schematic diagram showing that a sound acquiring device which collects the strongest sound signal is determined as the sound acquiring device meeting the preset condition according to an exemplary embodiment.

Fig. 2I is a flowchart showing a method for identifying a sound acquiring device closest to the surveilled object making sounds in surveillance images among all sound acquiring devices bound to the monitoring device according to an exemplary embodiment.

Fig. 2J is a schematic diagram showing that the monitoring device determines a sound acquiring device for monitoring the sounds from the surveilled object making sounds according to an exemplary embodiment.

Fig. 2K is a flowchart showing a method for controlling a sound acquiring device to collect sounds from the surveilled object according to an exemplary embodiment.

Fig. 2L is a schematic diagram showing that a monitoring device controls a mobile device to move to a location of a surveilled object making sounds in survalliance images according to an exemplary embodiment.

Fig. 3 is a block diagram showing a device for collecting sounds corresponding to surveillance images according to an exemplary embodiment.

Fig. 4 is a block diagram showing a device for collecting sounds corresponding to surveillance images according to another exemplary embodiment.

Fig. 5 is a block diagram showing a device for collecting sounds corresponding to surveillance images according to an exemplary embodiment.

DETAILED DESCRIPTION

[0027] Reference will now be made in detail to exemplary embodiments, examples of which are illustrated in the accompanying drawings. The following description refers to the accompanying drawings in which the same numbers in different drawings represent the same or similar elements unless otherwise represented. The implementations set forth in the following description of exemplary embodiments do not represent all implementations consistent with the present invention. Instead, they are merely examples of devices and methods consistent with aspects related to the present invention as recited in the appended claims.

[0028] Fig. 1 is a flowchart showing a method for collecting sounds corresponding to surveillance images according to an exemplary embodiment. As shown in Fig. 1, the method for collecting sounds corresponding to surveillance images includes the following steps.

[0029] In step 101, when surveillance images are collected by a camera of a monitoring device, whether there is a surveilled object making sounds in the surveillance images is detected.

[0030] Generally, a monitoring device can collect not only surveillance images but also sounds made by a surveilled object in the surveillance images.

[0031] In step 102, if there is a surveilled object making sounds in the surveillance images, a sound acquiring device meeting a preset condition is determined. The preset condition requires that a signal strength of the sounds collected by the sound acquiring device from the surveilled object is greater than a signal strength of the sounds collected by the monitoring device from the surveilled object.

[0032] At present, even though a monitoring device can clearly collect surveillance images, the monitoring device cannot effectively collect the sounds from a surveilled object in the surveillance images. For example, if the surveilled object in the surveillance images makes very weak sounds or the surveilled object in the surveillance images is relatively far from the monitoring device, the sounds from the surveilled object cannot be clearly collected by the monitoring device, and may even be automatically filtered by the monitoring device. Thus, an additional sound acquiring device meeting a preset condition is required to collect the sounds from the surveilled object.

[0033] The preset condition here may define that a signal strength of the sounds collected by the sound acquiring device from the surveilled object is greater than a signal strength of the sounds collected by the monitoring device from the surveilled object. When the preset condition is met, the sound acquiring device is turned on to collect the sounds from the surveilled object.

[0034] The sound acquiring device here may be devices having sound collecting members such as a mobile phone, an audio equipment, or a Bluetooth earphone.

[0035] In step 103, the sound acquiring device is con-

trolled to collect the sounds from the surveilled object.

[0036] When the monitoring device detects that the surveilled object in the surveillance images makes sounds but the sounds from the surveilled object are too weak or the surveilled object is too far from the monitoring device, the monitoring device cannot clearly collect the sounds from the surveilled object. At this time, the monitoring device may control the sound acquiring device meeting the preset condition to collect the sounds from the surveilled object.

[0037] In view of the above, in the method for collecting sounds corresponding to surveillance images provided by embodiments of the present invention, a surveilled object making sounds in the surveillance images is identified, and a sound acquiring device which collects sounds of a greater signal strength from the surveilled object than the signal strength of the sounds collected by the monitoring device from the surveilled object is controlled to collect the sounds from the surveilled object. Thus, the present invention can solve the problem related to a monitoring device not being able to clearly collect sounds from a surveilled object which is relatively far from the monitoring device. The monitoring device controls the sound acquiring device to collect the sounds from the surveilled object, and thus the surveillance information becomes more complete.

[0038] In a possible implementation, the step 102 in Fig. 1 may be replaced by steps 102a to 102b. Fig. 2A is a flowchart showing detection of whether there is a surveilled object making sounds in the surveillance images according to an exemplary embodiment. As shown in Fig. 2A, the detection of whether there is a surveilled object making sounds in surveillance images by the monitoring device may include the following steps.

[0039] In step 102a, whether sounds corresponding to the surveillance images collected by the monitoring device are smaller than a preset sound threshold is detected.

[0040] When the sounds made by the surveilled object in the surveillance images are too weak or the surveilled object in the surveillance images is too far from the monitoring device, the sounds made by the surveilled object cannot be clearly collected by the monitoring device. Thus, the preset sound threshold here refers to a threshold determining whether the monitoring device can clearly collect the sounds corresponding to the surveillance images. When the sounds corresponding to the surveillance images collected by the monitoring device are smaller than the preset sound threshold, it is indicated that the monitoring device cannot collect the sounds corresponding to the surveillance images temporarily.

[0041] In step 102b, if the sounds corresponding to the surveillance images collected by the monitoring device are smaller than the preset sound threshold, the sound acquiring device meeting the preset condition is determined.

[0042] When the sounds corresponding to the surveillance images collected by the monitoring device are smaller than the preset sound threshold, it is indicated that the monitoring device cannot collect the sounds corresponding to the surveillance images temporarily. Thus, the sound acquiring device meeting the preset condition needs to be determined to collect the sounds corresponding to the surveillance images instead of the monitoring device.

[0043] In view of the above, in the method for collecting sounds corresponding to surveillance images provided by embodiments of the present invention, a surveilled object making sounds in the surveillance images is identified, and a sound acquiring device which collects sounds of a greater signal strength from the surveilled object than the signal strength of the sounds collected by the monitoring device from the surveilled object is controlled to collect the sounds from the surveilled object. Thus, the present invention can solve the problem relating to a monitoring device not being able to clearly collect sounds from a surveilled object which is relatively far from the monitoring device. The monitoring device controls the sound acquiring device to collect the sounds from the surveilled object, and thus the surveillance information becomes more complete.

[0044] In addition, when the monitoring device cannot clearly collect the sounds made by the surveilled object, the monitoring device may send a collection instruction to a sound acquiring device meeting a preset condition to instruct the sound acquiring device to collect the sounds from the surveilled object. The preset condition requires that the sound acquiring device is closest to the surveilled object, or that the sound acquiring device collects the strongest sound signal, or that the sound acquiring device is a mobile device bound to the monitoring device.

[0045] A possible implementation is shown in Fig. 2B which is a flowchart showing a method for detecting whether there is a surveilled object making sounds in the surveillance images according to an exemplary embodiment. As shown in Fig. 2B, step 101 may include the following three implementations.

[0046] In a first implementation 101a, whether there is an image changing element making sounds in the surveillance images is detected. The image changing element includes a surveilled object entering into the surveillance images or a surveilled object whose lips move.

[0047] The image changing element here may refer to a person or an object moving into the surveillance images, or lips movement of a person or an animal.

[0048] Fig. 2C is a schematic diagram showing surveilled objects entering into the surveillance images according to an exemplary embodiment. As shown in Fig. 2C, when a person 11, a cat 12 and a dog 13 move into the images surveilled by a monitoring device 10, the monitoring device 10 can deem the person 11, the cat 12 and the dog 13 as the surveilled objects. Fig. 2D is a schematic diagram showing surveilled objects whose lips move in the surveillance images according to an exemplary embodiment. As shown in Fig. 2D, when persons

11a and 11b in the images surveilled by the monitoring device 10 are talking, the monitoring device 10 can deem the persons 11a and 11b as the surveilled objects when detecting that the lips of the persons 11a and 11b move.

[0049] In a second implementation 101b, whether an object of a designated type in the surveillance images makes sounds is detected. The object of the designated type includes at least one of a door, a window and a home appliance.

[0050] Sometimes, when a door, a window and a home appliance in surveillance images make sounds (for example when a person is knocking at a door, a TV is turned on, and a fan runs), no significant changes will occur in the surveillance images. Thus, the object of the designated type refers to an object which does not move and the camera needs to determine whether to acquire sounds of the object according to the collected sounds.

[0051] Fig. 2E is a schematic diagram showing an object of a designated type making sounds in the surveillance images according to an exemplary embodiment. As shown in Fig. 2E, when a door 14 is knocked at, a TV15 is turned on and a fan 16 runs in the surveillance images, even though there is no significant changes in the whole surveillance images, the monitoring device 10 may still deem the door 14, the TV15 and the fan 16 as the surveilled objects according to the knock, the sound made by TV program and the sound made by blades of the fan.

[0052] In a third implementation 101c, according to a designated index reported by a detection device, whether there is a surveilled object making sounds in the surveillance images is detected. The detection device is a device bound to the monitoring device for detecting the designated index, and the designated index includes one of on/off states, air quality parameters, temperature, parameters indicating changes of heat sources or physiological feature parameters

[0053] When the sounds made by the surveilled object in the surveillance images are too weak, the monitoring device may determine which surveilled objects in the surveillance images make sounds via a detection device. The detection device here refers to a device for detecting states of the surveilled objects which is installed around the surveilled objects or on the surveilled objects.

[0054] For example, the detection device is installed on a switch of a home appliance. When the home appliance is turned on, the detection device sends the turning-on state data back to the monitoring device. After receiving the data, the monitoring device determines that the home appliance is the surveilled object, and sends a collection instruction to a sound acquiring device meeting a preset condition to instruct the sound acquiring device meeting the preset condition to collect the sounds from the home appliance. In another example, the detection device may be installed nearby a window or a door. When the door or the window is open, and the detection device senses the temperature changes of the ambient air, the detection device feeds the temperature change data back to the monitoring device. After receiving the data, the monitoring device determines that the window or the door is the surveilled object, and sends a collection instruction to a sound acquiring device meeting a preset condition to instruct the sound acquiring device meeting the preset condition to collect the sounds from the door or the window.

[0055] Fig. 2F is a schematic diagram showing detection of whether there is a surveilled object making sounds in the surveillance images according to a designated index reported by a detection device according to an exemplary embodiment. As shown in Fig. 2F, when a door 18 in the surveillance images is open, a sensor 17 installed above the door 18 senses the temperature changes of the ambient air, and feeds the temperature change data back to a monitoring device 10 which receives the feedback data and sends a collection instruction to a sound acquiring device meeting a preset condition to instruct the sound acquiring device meeting the preset condition to collect the sounds from the door 18.

[0056] A possible implementation is shown in Fig. 2G which is a flowchart showing a method for determining a sound acquiring device meeting a preset condition according to an exemplary embodiment. As shown in Fig. 2G, the determination of the sound acquiring device meeting a preset condition (step 102b) may include the following three implementations:

> In a first implementation 102b1, a sound acquiring device closest to the surveilled object making sounds is identified among all sound acquiring devices bound to the monitoring device, and the identified sound acquiring device is determined as the sound acquiring device meeting the preset condition.

[0057] When the preset condition requires that the sound acquiring device is closest to the surveilled object, the monitoring device may identify a sound acquiring device closest to the surveilled object making sounds according to the surveillance images, and determine the sound acquiring device as the sound acquiring device meeting the preset condition.

[0058] In a second implementation 102b2, all sound acquiring devices bound to the monitoring device are controlled to collect the sounds from the surveilled object making sounds, and a sound acquiring device which collects a strongest sound signal is determined as the sound acquiring device meeting the preset condition.

[0059] When the preset condition requires that the sound acquiring device collects the strongest sound signal, the monitoring device may send a collection instruction to all sound acquiring devices to instruct all of the sound acquiring devices to collect the sounds from the surveilled object in the surveillance images. After receiving the collection instruction, all the sound acquiring devices start to collect the sounds from the surveilled object and send the collected sounds to the monitoring device for testing. The monitoring device receives and tests all

the sounds collected by the sound acquiring devices, determines a sound acquiring device which collects the strongest sound signal and determines the sound acquiring device as the sound acquiring device meeting the preset condition.

[0060] Fig. 2H is a schematic diagram showing determination of the sound acquiring device which collects the strongest sound signal as the sound acquiring device meeting the preset condition according to an exemplary embodiment. As shown in Fig. 2H, there are three sound acquiring devices in surveillance images, i.e., an audio equipment 19, a Bluetooth earphone 20 and a mobile phone 21. If a person 11 enters into the surveillance images, a monitoring device 10 determines the person 11 as the surveilled object, and then sends a collection instruction to the audio equipment 19, the Bluetooth earphone 20 and the mobile phone 21. After receiving the collection instruction, the three sound acquiring devices collect sounds from the person 11 and send the collected sounds to the monitoring device 10. The monitoring device 10 tests the sounds collected by the three sound acquiring devices, and determines that the Bluetooth earphone 20 collects the strongest sound signal, and thereby determines the Bluetooth earphone 20 as the sound acquiring device meeting the preset condition.

[0061] In a third implementation 102b3, at least one mobile device bound to the monitoring device is determined as the sound acquiring device meeting the preset condition.

[0062] The monitoring device may be bound to mobile devices which move in house interior, for example, a mobile terminal, a home cleaning robot, a remotely controlled plane. When the surveilled object is too far from the monitoring device, these mobile devices are determined as the sound acquiring devices meeting the preset condition.

[0063] A possible implementation is shown in Fig. 2I which is a flowchart showing a method for identifying a sound acquiring device closest to the surveilled object making sounds in surveillance images among all sound acquiring devices bound to the monitoring device according to an exemplary embodiment. As shown in Fig. 2I, the first implementation 102b1 in step 102b may include the following two implementations:

> In a first implementation 102b1a, by an indoor positioning technology, the sound acquiring device closest to the surveilled object making sounds is identified among all sound acquiring devices bound to the monitoring device.

[0064] Monitoring device may be installed in a plurality of rooms, and the monitoring devices may be connected via wireless or wired networks to form an indoor positioning system. Thus, persons, animals or home appliances and the like in the interior of the rooms may be surveilled. When a surveilled object making sounds appears in surveillance images of a monitoring device, a sound acquiring device closest to the surveilled object can be identified from all the sound acquiring devices bound to the monitoring device, and the monitoring device can send a collection instruction to the sound acquiring device to instruct the sound acquiring device to collect the sounds from the surveilled object.

[0065] In a second implementation 102b1b, by an image recognition technology, a location of the surveilled object making sounds is determined, and a sound acquiring device for monitoring the sounds from the surveilled object making sounds is determined according to respective surveillance area ranges corresponding to the sound acquiring devices.

[0066] Each sound acquiring device has individual surveillance area range. The surveillance range data of all sound acquiring devices may be recorded in the monitoring device. According to the surveillance images, the monitoring device can determine a surveillance area range of a sound acquiring device in which the surveilled object making sounds is, and thus the monitoring device can determine the sound acquiring device for monitoring the surveilled object.

[0067] Fig. 2J is a schematic diagram showing that the monitoring device determines a sound acquiring device for monitoring the sounds from the surveilled object making sounds according to an exemplary embodiment. As shown in Fig. 2J, there are two corners in the current surveillance images where monitoring devices 10a and 10b are installed. The monitoring devices 10a and 10b are connected with each other via wireless or wired networks to form an indoor positioning system. When a TV 15 in the surveillance images of the monitoring device 10a is turned on, according to the surveillance images, the monitoring device 10a determines that there is only one sound acquiring device (i.e., an audio equipment 19) in the surveillance images, but the TV 15 is not in the surveillance area range 19a of the audio equipment 19, and thus the monitoring device 10a sends an inquiry instruction to the monitoring device 10b to inquire whether there is a mobile device bound to all the monitoring devices in the surveillance images of the monitoring device 10b. After receiving the inquiry instruction, the monitoring device 10b determines that a home cleaning robot 23 in corresponding surveillance images is bound to all the monitoring devices, and sends a moving instruction to the home cleaning robot 23 to instruct the home cleaning robot 23 to move into the surveillance images of the monitoring device 10a. The monitoring device 10a sends the home cleaning robot 23 a moving instruction for instructing the home cleaning robot 23 to move to the location where the TV 15 is to make the TV 15 in the surveillance area range 23a of the home cleaning robot 23, and a collection instruction for instructing the home cleaning robot 23 to collect the sounds from the TV 15.

[0068] Still referring to Fig. 2J, when a dog 13 enters into the surveillance images of the monitoring device 10b, according to the surveillance images, the monitoring device 10b determines that there are two sound acquiring

devices (i.e., a mobile phone 21 and the home cleaning robot 23) in the surveillance images, the dog 13 is in the surveillance area range 21a of the mobile phone 21 but not in the surveillance area range 23a of the home cleaning robot 23. At this time, the monitoring device 10b may send a collection instruction to the mobile phone 21 to instruct to the mobile phone 21 to collect the sounds from the dog 13. Further, in order to arrive at better sound collection effects, the monitoring device 10b may send the home cleaning robot 23 a moving instruction for instructing the home cleaning robot 23 to move to a location nearby the dog 13 to make the dog 13 in the surveillance area range 23a of the home cleaning robot 23, and a collection instruction for instructing the home cleaning robot 23 to collect the sounds from the dog 13.

[0069] A possible implementation is shown in Fig. 2K which is a flowchart showing a method for controlling a sound acquiring device to collect sounds from a surveilled object according to an exemplary embodiment. As shown in Fig. 2K, step 103 may include the following three implementations:

> In a first implementation 103a, when the sound acquiring device meeting the preset condition is a sound acquiring device closest to the surveilled object making sounds, the sound acquiring device is controlled to collect sounds.

[0070] When the monitoring device cannot collect the sounds corresponding to the surveilled object making sounds in the surveillance images temporarily, the monitoring device can send a collection instruction to a sound acquiring device closest to the surveilled object making sounds to instruct the sound acquiring device to collect the sounds made by the surveilled object.

[0071] In a second implementation 103b, when the sound acquiring device meeting the preset condition is a sound acquiring device which collects the strongest sound signal, the sound acquiring device is controlled to collect sounds.

[0072] When the monitoring device cannot collect the sounds corresponding to the surveilled object making sounds in the surveillance images temporarily, the monitoring device can send a collection instruction to a sound acquiring device which collects the strongest sound signal to instruct the sound acquiring device to collect the sounds made by the surveilled object.

[0073] In a third implementation 103c, when the sound acquiring device meeting the preset condition is a mobile device bound to the monitoring device, the mobile device is controlled to move to a location of the surveilled object making sounds in the surveillance images, and the mobile device is controlled to collect sounds at the location where the surveilled object is.

[0074] When the monitoring device determines that the sound acquiring device meeting the preset condition is a mobile device bound to the monitoring device, the monitoring device may send the mobile device a moving instruction for instructing the mobile device to move to a location nearby the surveilled object making sounds in the surveillance images, and a collection instruction for instructing the mobile device to collect the sounds from the surveilled object which is making sounds.

[0075] Fig. 2L is a schematic diagram showing that a monitoring device controls a mobile device to move to a location of a surveilled object making sounds in surveillance images according to an exemplary embodiment. As shown in Fig. 2L, when a monitoring device 10 cannot detect the sounds made by a cat 12, the monitoring device 10 determines the sound acquiring devices meeting the preset condition as a remotely controlled plane 22 and a home cleaning robot 23 bound to the monitoring device 10, and then sends the remotely controlled plane 22 or the home cleaning robot 23 a moving instruction and a collection instruction. The moving instruction instructs the remotely controlled plane 22 or the home cleaning robot 23 to move to the location where the cat 12 is, and the collection instruction instructs the remotely controlled plane 22 or the home cleaning robot 23 to collect the sounds made by the cat 12.

[0076] The following are embodiments of the device of the present invention, which may be used to perform the embodiments of the method of the present invention. For the undisclosed details in the embodiments of the device, refer to the embodiments of the method of the present invention.

[0077] Fig. 3 is a block diagram showing a device for collecting sounds corresponding to surveillance images according to an exemplary embodiment. As shown in Fig. 3, the device for collecting sounds corresponding to surveillance images includes but not limited to a detection module 201, a determination module 202 and a control module 203.

[0078] The detection module 201 is configured to, when surveillance images are collected by a camera of a monitoring device, detect whether there is a surveilled object making sounds in the surveillance images.

[0079] The determination module 202 is configured to, if the detection module 201 detects that there is a surveilled object making sounds in the surveillance images, determine a sound acquiring device meeting a preset condition which requires that a signal strength of the sounds collected by the sound acquiring device from the surveilled object is greater than a signal strength of the sounds collected by the monitoring device from the surveilled object.

[0080] The control module 203 is configured to control the sound acquiring device determined by the determination module 202 to collect the sounds from the surveilled object.

[0081] A possible implementation is shown in Fig. 4 which is a block diagram showing a device for collecting sounds corresponding to surveillance images according to another exemplary embodiment. As shown in Fig. 4, the detection module 201 includes a first detection submodule 201a, a second detection submodule 201b, and

a third detection submodule 201c.

[0082] The first detection submodule 201a is configured to detect whether there is an image changing element making sounds in the surveillance images. The image changing element includes a surveilled object entering into the surveillance images or a surveilled object whose lips move.

[0083] The second detection submodule 201b is configured to detect whether an object of a designated type in the surveillance images makes sounds. The object of the designated type includes at least one of a door, a window and a home appliance.

[0084] The third detection submodule 201c is configured to, according to a designated index reported by a detection device, detect whether there is a surveilled object making sounds in the surveillance images, wherein the detection device is a device bound to the monitoring device for detecting the designated index, and the designated index includes one of on/off states, air quality parameters, temperature, parameters indicating changes of heat sources or physiological feature parameters.

[0085] In a possible implementation, still referring to Fig. 4, the determination module 202 includes a fourth detection submodule 202a and a first determination submodule 202b.

[0086] The fourth detection submodule 202a is configured to detect whether sounds corresponding to the surveillance images collected by the monitoring device are smaller than a preset sound threshold.

[0087] The first determination submodule 202b is configured to, if the fourth detection submodule 202a detects that the sounds corresponding to the surveillance images collected by the monitoring device are smaller than the preset sound threshold, determine the sound acquiring device meeting the preset condition.

[0088] In a possible implementation, still referring to Fig. 4, the first determination submodule 202b includes a first identifying submodule 202b1, a second identifying submodule 202b2, or a second determination submodule 202b3.

[0089] The first identifying submodule 202b1 is configured to identify a sound acquiring device closest to the surveilled object making sounds among all sound acquiring devices bound to the monitoring device, and determine the identified sound acquiring device as the sound acquiring device meeting the preset condition.

[0090] The second identifying submodule 202b2 is configured to control all sound acquiring devices bound to the monitoring device to collect the sounds from the surveilled object making sounds, and determine a sound acquiring device which collects a strongest sound signal as the sound acquiring device meeting the preset condition.

[0091] The second determination submodule 202b3 is configured to determine at least one mobile device bound to the monitoring device as the sound acquiring device meeting the preset condition.

[0092] In a possible implementation, still referring to Fig. 4, the first identifying submodule 202b1 includes a third identifying submodule 202b1a or a fourth determination submodule 202b1b.

[0093] The third identifying submodule 202b1a is configured to, by an indoor positioning technology, identify the sound acquiring device closest to the surveilled object making sounds among all sound acquiring devices bound to the monitoring device.

[0094] The fourth determination submodule 202b1b is configured to, by an image recognition technology, determine a location where the surveilled object making sounds is, and determine a sound acquiring device for monitoring the sounds from the surveilled object making sounds according to respective surveillance area ranges corresponding to the sound acquiring devices.

[0095] In a possible implementation, still referring to Fig. 4, the control module 203 includes a first control submodule 203a, a second control submodule 203b or a third control submodule 203c.

[0096] The first control submodule 203a is configured to, when the sound acquiring device meeting the preset condition is a sound acquiring device closest to the surveilled object making sounds, control the sound acquiring device to collect sounds.

[0097] The second control submodule 203b is configured to, when the sound acquiring device meeting the preset condition is a sound acquiring device which collects the strongest sound signal, control the sound acquiring device to collect sounds.

[0098] The third control submodule 203c is configured to, when the sound acquiring device meeting the preset condition is a mobile device bound to the monitoring device, control the mobile device to move to a location of the surveilled object making sounds in the surveillance images, and control the mobile device to collect sounds at the location where the surveilled object is.

[0099] In view of the above, in the device for collecting sounds corresponding to surveillance images provided by embodiments of the present invention, a surveilled object making sounds in the surveillance images is identified, and a sound acquiring device which collects sounds of a greater signal strength from the surveilled object than the signal strength of the sounds collected by the monitoring device from the surveilled object is controlled to collect the sounds from the surveilled object. Thus, the present invention can solve the problem related to a monitoring device not being able to clearly collect sounds from a surveilled object which is relatively far from the monitoring device. The monitoring device controls the sound acquiring device to collect the sounds from the surveilled object, and thus the surveillance information becomes more complete.

[0100] In addition, when the monitoring device cannot clearly collect the sounds made by the surveilled object, the monitoring device may send a collection instruction to a sound acquiring device meeting a preset condition to instruct the sound acquiring device to collect the sounds from the surveilled object. The preset condition

requires that the sound acquiring device is closest to the surveilled object, or that the sound acquiring device collects the strongest sound signal, or that the sound acquiring device is a mobile device bound to the monitoring device.

[0101] With respect to the devices in the above embodiments, the specific manners for performing operations for individual modules therein have been described in detail in the embodiments regarding the methods, which will not be elaborated herein.

[0102] An exemplary embodiment of the present invention provides a device for collecting sounds corresponding to surveillance images, which is capable of realizing the method for collecting sounds corresponding to surveillance images provided by the present invention. The device for collecting sounds corresponding to surveillance images includes:

a processor; and
a memory for storing instructions executable by the processor;
wherein the processor is configured to:

when surveillance images are collected by a camera of a monitoring device, detect whether there is a surveilled object making sounds in the surveillance images;
if there is a surveilled object making sounds in the surveillance images, determine a sound acquiring device meeting a preset condition which requires that a signal strength of the sounds collected by the sound acquiring device from the surveilled object is greater than a signal strength of the sounds collected by the monitoring device from the surveilled object; and
control the sound acquiring device to collect the sounds from the surveilled object.

[0103] Fig. 5 is a block diagram showing a device for collecting sounds corresponding to surveillance images according to an exemplary embodiment. For example, the device 400 may be a mobile phone, a computer, a digital broadcast terminal, a messaging device, a gaming console, a tablet, a medical device, exercise equipment, a personal digital assistant, and the like.

[0104] Referring to Fig. 5, the device 400 may include one or more of the following components: a processing component 402, a memory 404, a power component 406, a multimedia component 408, an audio component 410, an input/output (I/O) interface 412, a sensor component 414, and a communication component 416.

[0105] The processing component 402 typically controls overall operations of the device 400, such as the operations associated with display, telephone calls, data communications, camera operations, and recording operations. The processing component 402 may include one or more processors 418 to execute instructions to perform all or part of the steps in the above described methods. Moreover, the processing component 402 may include one or more modules which facilitate the interaction between the processing component 402 and other components. For instance, the processing component 402 may include a multimedia module to facilitate the interaction between the multimedia component 408 and the processing component 402.

[0106] The memory 404 is configured to store various types of data to support the operation of the device 400. Examples of such data include instructions for any applications or methods operated on the device 400, contact data, phonebook data, messages, pictures, video, etc. The memory 404 may be implemented using any type of volatile or non-volatile memory devices, or a combination thereof, such as a static random access memory (SRAM), an electrically erasable programmable read-only memory (EEPROM), an erasable programmable read-only memory (EPROM), a programmable read-only memory (PROM), a read-only memory (ROM), a magnetic memory, a flash memory, a magnetic or optical disk.

[0107] The power component 406 provides power to various components of the device 400. The power component 406 may include a power management system, one or more power sources, and any other components associated with the generation, management, and distribution of power in the device 400.

[0108] The multimedia component 408 includes a screen providing an output interface between the device 400 and the user. In some embodiments, the screen may include a liquid crystal display (LCD) and a touch panel (TP). If the screen includes the touch panel, the screen may be implemented as a touch screen to receive input signals from the user. The touch panel includes one or more touch sensors to sense touches, swipes, and gestures on the touch panel. The touch sensors may not only sense a boundary of a touch or swipe action, but also sense a period of time and a pressure associated with the touch or swipe action. In some embodiments, the multimedia component 408 includes a front camera and/or a rear camera. The front camera and the rear camera may receive an external multimedia datum while the device 400 is in an operation mode, such as a photographing mode or a video mode. Each of the front camera and the rear camera may be a fixed optical lens system or have focus and optical zoom capability.

[0109] The audio component 410 is configured to output and/or input audio signals. For example, the audio component 410 includes a microphone ("MIC") configured to receive an external audio signal when the device 400 is in an operation mode, such as a call mode, a recording mode, and a voice recognition mode. The received audio signal may be further stored in the memory 404 or transmitted via the communication component 416. In some embodiments, the audio component 410 further includes a speaker to output audio signals.

[0110] The I/O interface 412 provides an interface between the processing component 402 and peripheral interface modules, such as a keyboard, a click wheel, but-

tons, and the like. The buttons may include, but are not limited to, a home button, a volume button, a starting button, and a locking button.

[0111] The sensor component 414 includes one or more sensors to provide status assessments of various aspects of the device 400. For instance, the sensor component 414 may detect an open/closed status of the device 400, relative positioning of components, e.g., the display and the keypad, of the device 400, a change in position of the device 400 or a component of the device 400, a presence or absence of user contact with the device 400, an orientation or an acceleration/deceleration of the device 400, and a change in temperature of the device 400. The sensor component 414 may include a proximity sensor configured to detect the presence of nearby objects without any physical contact. The sensor component 414 may also include a light sensor, such as a CMOS or CCD image sensor, for use in imaging applications. In some embodiments, the sensor component 414 may also include an accelerometer sensor, a gyroscope sensor, a magnetic sensor, a pressure sensor, or a temperature sensor.

[0112] The communication component 416 is configured to facilitate communication, wired or wirelessly, between the device 400 and other devices. The device 400 can access a wireless network based on a communication standard, such as WiFi, 2G, or 3G, or a combination thereof. In one exemplary embodiment, the communication component 416 receives a broadcast signal or broadcast associated information from an external broadcast management system via a broadcast channel. In one exemplary embodiment, the communication component 416 further includes a near field communication (NFC) module to facilitate short-range communications. For example, the NFC module may be implemented based on a radio frequency identification (RFID) technology, an infrared data association (IrDA) technology, an ultra-wideband (UWB) technology, a Bluetooth (BT) technology, and other technologies.

[0113] In exemplary embodiments, the device 400 may be implemented with one or more application specific integrated circuits (ASICs), digital signal processors (DSPs), digital signal processing devices (DSPDs), programmable logic devices (PLDs), field programmable gate arrays (FPGAs), controllers, micro-controllers, microprocessors, or other electronic components, for performing the above described methods for collecting sounds corresponding to surveillance images.

[0114] In exemplary embodiments, there is also provided a non-transitory computer-readable storage medium including instructions, such as included in the memory 404, executable by the processor 418 in the device 400, for performing the above-described methods for collecting sounds corresponding to surveillance images. For example, the non-transitory computer-readable storage medium may be a ROM, a RAM, a CD-ROM, a magnetic tape, a floppy disc, an optical data storage device, and the like.

[0115] Other embodiments of the present invention will be apparent to those skilled in the art from consideration of the specification and practice of the invention disclosed here. This application is intended to cover any variations, uses, or adaptations of the present invention following the general principles thereof and including such departures from the present invention as come within known or customary practice in the art. It is intended that the specification and examples be considered as exemplary only, with a true scope and spirit of the present invention being indicated by the following claims.

[0116] It will be appreciated that the present invention is not limited to the exact construction that has been described above and illustrated in the accompanying drawings, and that various modifications and changes can be made without departing from the scope thereof. It is intended that the scope of the present invention only be limited by the appended claims.

Claims

1. A method for collecting sounds corresponding to surveillance images, comprising:

   when surveillance images are collected by a camera of a monitoring device, detecting (101) whether there is a surveilled object making sounds in the surveillance images;
   if there is a surveilled object making sounds in the surveillance images, determining (102) a sound acquiring device meeting a preset condition which requires that a signal strength of the sounds collected by the sound acquiring device from the surveilled object is greater than a signal strength of the sounds collected by the monitoring device from the surveilled object; and
   controlling (103) the sound acquiring device to collect the sounds from the surveilled object.

2. The method according to claim 1, wherein detecting (101) whether there is a surveilled object making sounds in the surveillance images, comprises:

   detecting (101a) whether there is an image changing element making sounds in the surveillance images, wherein the image changing element comprises a surveilled object entering into the surveillance images or a surveilled object whose lips move; or
   detecting (101b) whether an object of a designated type in the surveillance images makes sounds, wherein the object of the designated type comprises at least one of a door, a window and a home appliance; or
   according to a designated index reported by a detection device, detecting (101c) whether there is a surveilled object making sounds in the sur-

veillance images, wherein the detection device is a device bound to the monitoring device for detecting the designated index, and the designated index comprises one of on/off states, air quality parameters, temperature, parameters indicating changes of heat sources or physiological feature parameters.

3. The method according to claim 1, wherein determining (102) a sound acquiring device meeting a preset condition comprises:

detecting (102a) whether sounds corresponding to the surveillance images collected by the monitoring device are smaller than a preset sound threshold; and
if the sounds corresponding to the surveillance images collected by the monitoring device are smaller than the preset sound threshold, determining (102b) the sound acquiring device meeting the preset condition.

4. The method according to claim 3, wherein determining (102) the sound acquiring device meeting the preset condition comprises:

identifying (102b1) a sound acquiring device closest to the surveilled object making sounds among all sound acquiring devices bound to the monitoring device, and determining the identified sound acquiring device as the sound acquiring device meeting the preset condition; or
controlling (102b2) all sound acquiring devices bound to the monitoring device to collect the sounds from the surveilled object making sounds, and determining a sound acquiring device which collects a strongest sound signal as the sound acquiring device meeting the preset condition; or
determining (102b3) at least one mobile device bound to the monitoring device as the sound acquiring device meeting the preset condition.

5. The method according to claim 4, wherein identifying (102b1) a sound acquiring device closest to the surveilled object making sounds among all sound acquiring devices bound to the monitoring device, comprises:

by an indoor positioning technology, identifying (102b1a) the sound acquiring device closest to the surveilled object making sounds among all sound acquiring devices bound to the monitoring device; or
by an image recognition technology, determining (102b1b) a location where the surveilled object making sounds is, and determining a sound acquiring device for monitoring the sounds from the surveilled object making sounds according to respective surveillance area ranges corresponding to the sound acquiring devices.

6. The method according to any one of claims 1 to 5, wherein controlling (103) the sound acquiring device to collect the sounds from the surveilled object, comprises:

when the sound acquiring device meeting the preset condition is a sound acquiring device closest to the surveilled object making sounds, controlling (103a) the sound acquiring device to collect sounds; or
when the sound acquiring device meeting the preset condition is a sound acquiring device which collects the strongest sound signal, controlling (103b) the sound acquiring device to collect sounds; or
when the sound acquiring device meeting the preset condition is a mobile device bound to the monitoring device, controlling (103c) the mobile device to move to a location of the surveilled object making sounds in the surveillance images, and controlling the mobile device to collect sounds at the location where the surveilled object is.

7. A device for collecting sounds corresponding to surveillance images, comprising:

a detection module (201) configured to, when surveillance images are collected by a camera of a monitoring device, detect whether there is a surveilled object making sounds in the surveillance images;
a determination module (202) configured to, if the detection module detects that there is a surveilled object making sounds in the surveillance images, determine a sound acquiring device meeting a preset condition which requires that a signal strength of the sounds collected by the sound acquiring device from the surveilled object is greater than a signal strength of the sounds collected by the monitoring device from the surveilled object; and
a control module (203) configured to control the sound acquiring device determined by the determination module to collect the sounds from the surveilled object.

8. The device according to claim 7, wherein the detection module (201) comprises:

a first detection submodule (201a) configured to detect whether there is an image changing element making sounds in the surveillance images, wherein the image changing element comprises

a surveilled object entering into the surveillance images or a surveilled object whose lips move; or

a second detection submodule (201b) configured to detect whether an object of a designated type in the surveillance images makes sounds, wherein the object of the designated type comprises at least one of a door, a window and a home appliance; or

a third detection submodule (201c) configured to, according to a designated index reported by a detection device, detect whether there is a surveilled object making sounds in the surveillance images, wherein the detection device is a device bound to the monitoring device for detecting the designated index, and the designated index comprises one of on/off states, air quality parameters, temperature, parameters indicating changes of heat sources or physiological feature parameters.

9. The device according to claim 7, wherein the determination module (202) comprises:

a fourth detection submodule (202a) configured to detect whether sounds corresponding to the surveillance images collected by the monitoring device are smaller than a preset sound threshold; and

a first determination submodule (202b) configured to, if the fourth detection submodule detects that the sounds corresponding to the surveillance images collected by the monitoring device are smaller than the preset sound threshold, determine the sound acquiring device meeting the preset condition.

10. The device according to claim 9, wherein the first determination submodule (202b) comprises:

a first identifying submodule (202b1) configured to identify a sound acquiring device closest to the surveilled object making sounds among all sound acquiring devices bound to the monitoring device, and determine the identified sound acquiring device as the sound acquiring device meeting the preset condition; or

a second identifying submodule (202b2) configured to control all sound acquiring devices bound to the monitoring device to collect the sounds from the surveilled object making sounds, and determine a sound acquiring device which collects a strongest sound signal as the sound acquiring device meeting the preset condition; or

a second determination submodule (202b3) configured to determine at least one mobile device bound to the monitoring device as the sound acquiring device meeting the preset condition.

11. The device according to claim 10, wherein the first identifying submodule (202b1) comprises:

a third identifying submodule (202b1a) configured to, by an indoor positioning technology, identify the sound acquiring device closest to the surveilled object making sounds among all sound acquiring devices bound to the monitoring device; or

a fourth determination submodule (202b1b) configured to, by an image recognition technology, determine a location where the surveilled object making sounds is, and determine a sound acquiring device for monitoring the sounds from the surveilled object making sounds according to respective surveillance area ranges corresponding to the sound acquiring devices.

12. The device according to any one of claims 7 to 11, wherein the control module (203) comprises:

a first control submodule (203a) configured to, when the sound acquiring device meeting the preset condition is a sound acquiring device closest to the surveilled object making sounds, control the sound acquiring device to collect sounds; or

a second control submodule (203b) configured to, when the sound acquiring device meeting the preset condition is a sound acquiring device which collects the strongest sound signal, control the sound acquiring device to collect sounds; or

a third control submodule (203c) configured to, when the sound acquiring device meeting the preset condition is a mobile device bound to the monitoring device, control the mobile device to move to a location of the surveilled object making sounds in the surveillance images, and control the mobile device to collect sounds at the location where the surveilled object is.

13. A device (400) for collecting sounds corresponding to surveillance images, comprising:

a processor (402); and

a memory (404) for storing instructions executable by the processor (402);

wherein the processor (402) is configured to:

when surveillance images are collected by a camera of a monitoring device, detect whether there is a surveilled object making sounds in the surveillance images;

if there is a surveilled object making sounds in the surveillance images, determine a

sound acquiring device meeting a preset condition which requires that a signal strength of the sounds collected by the sound acquiring device from the surveilled object is greater than a signal strength of the sounds collected by the monitoring device from the surveilled object; and
control the sound acquiring device to collect the sounds from the surveilled object.

14. A computer program which, when executed on a processor, performs a method according to any one of claims 1 to 6.

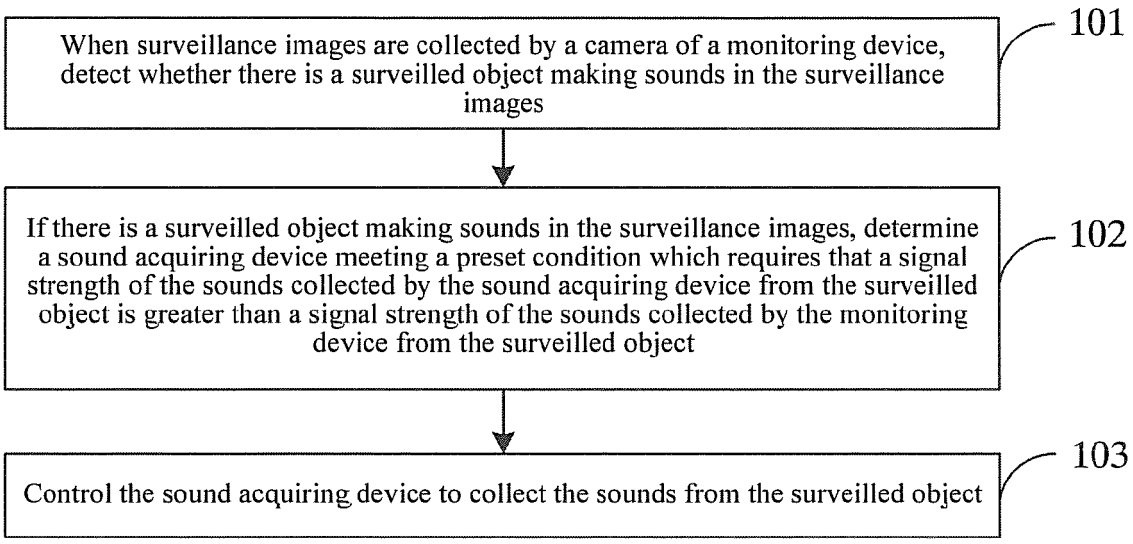

Fig. 1

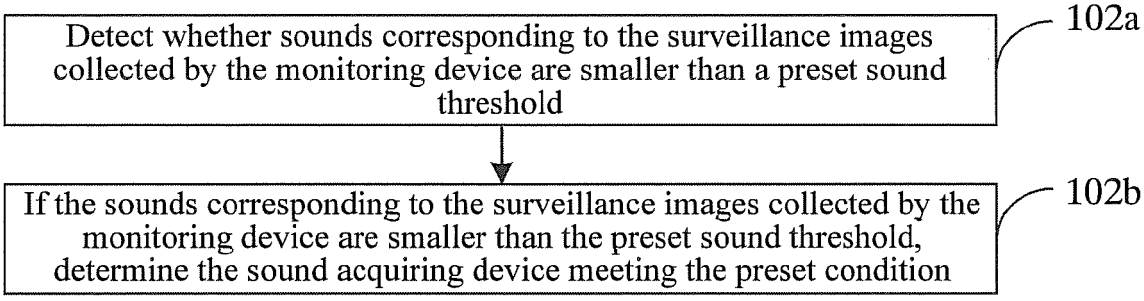

Fig. 2A

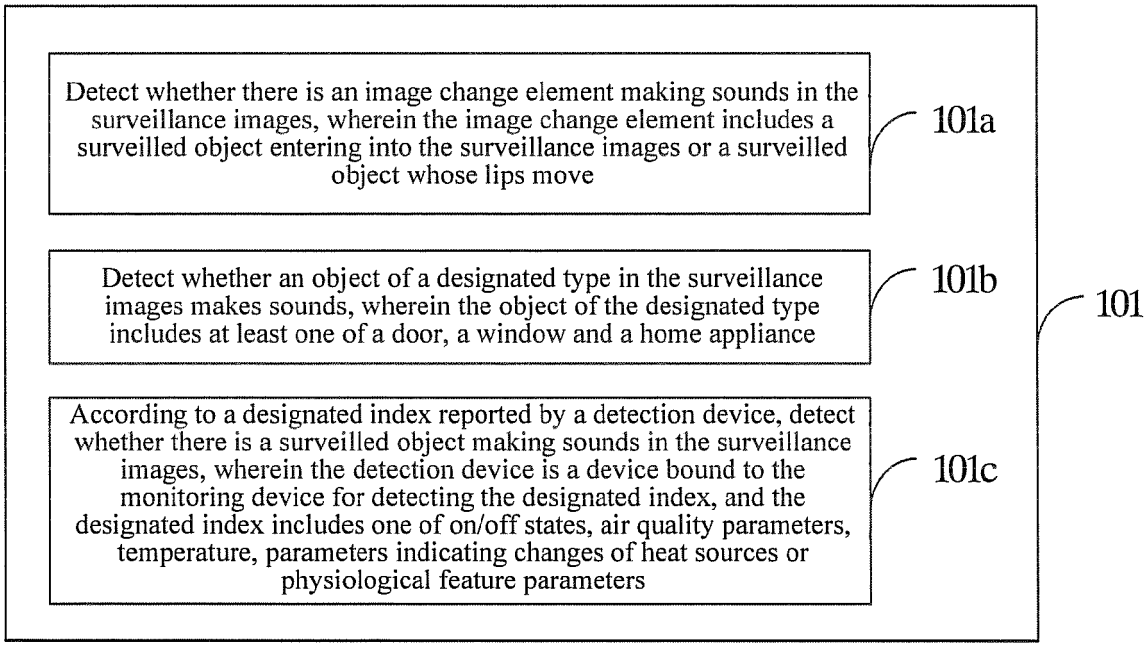

Fig. 2B

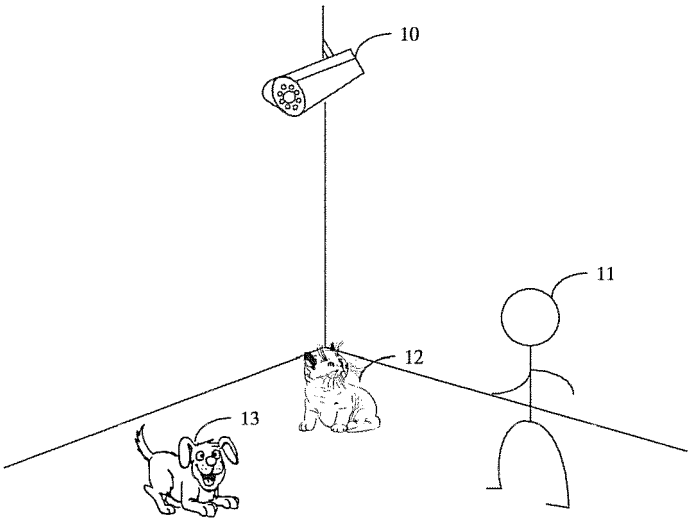

Fig. 2C

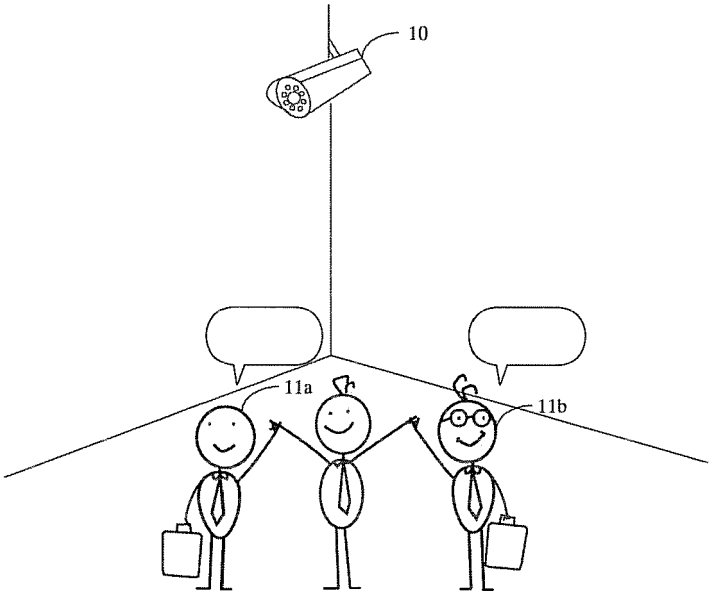

Fig. 2D

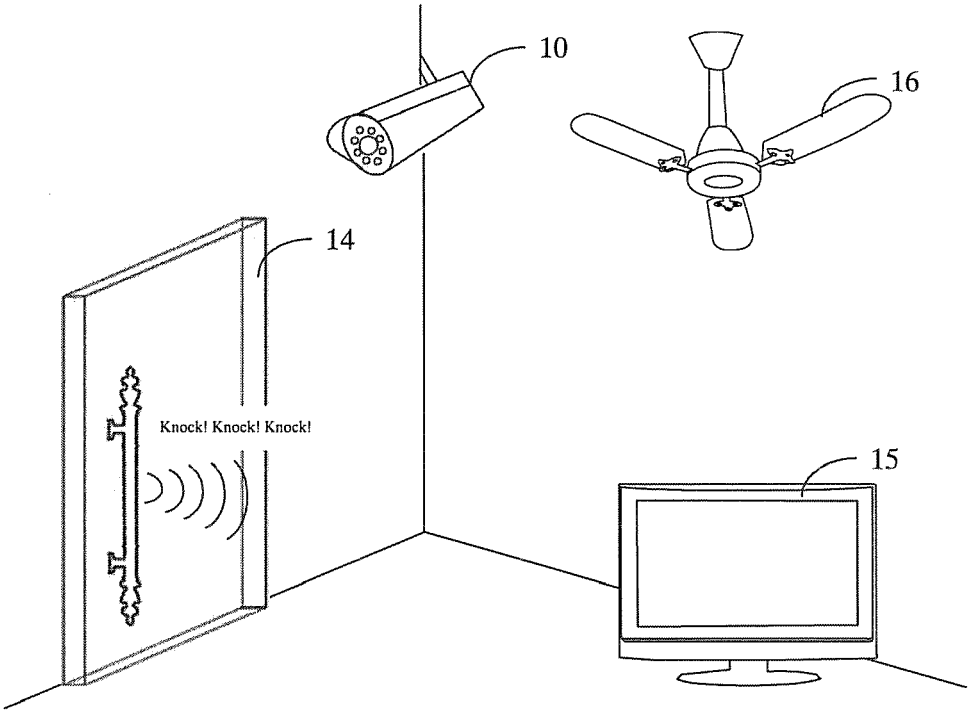

Fig. 2E

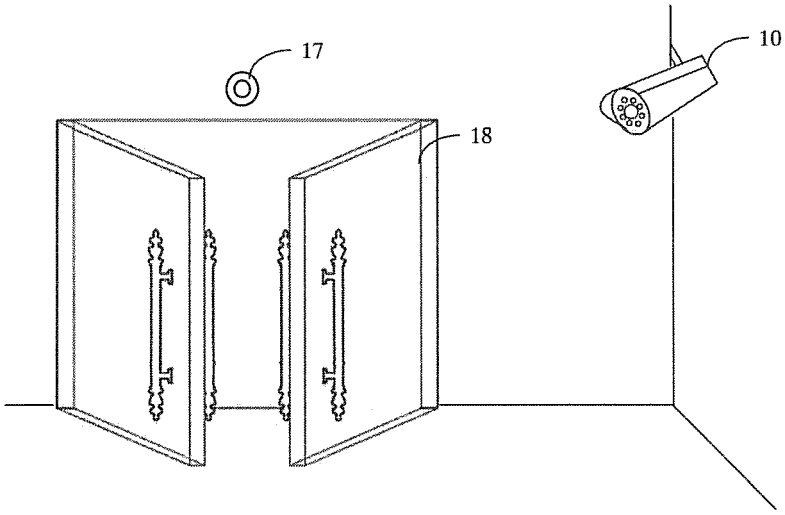

Fig. 2F

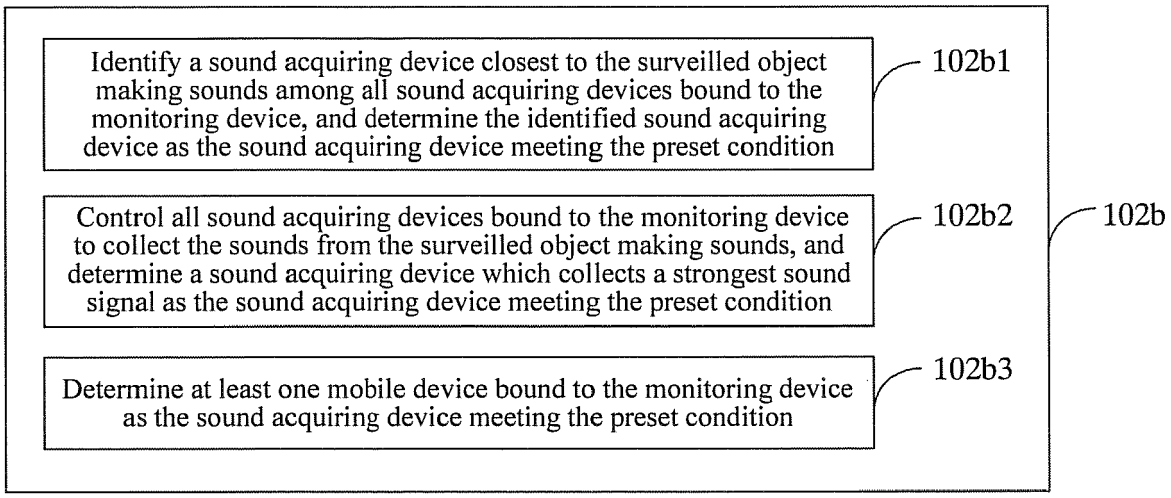

Fig. 2G

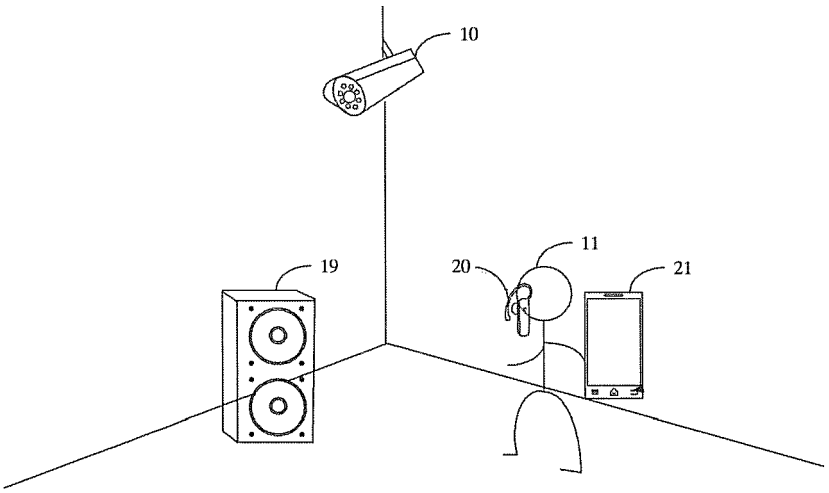

Fig. 2H

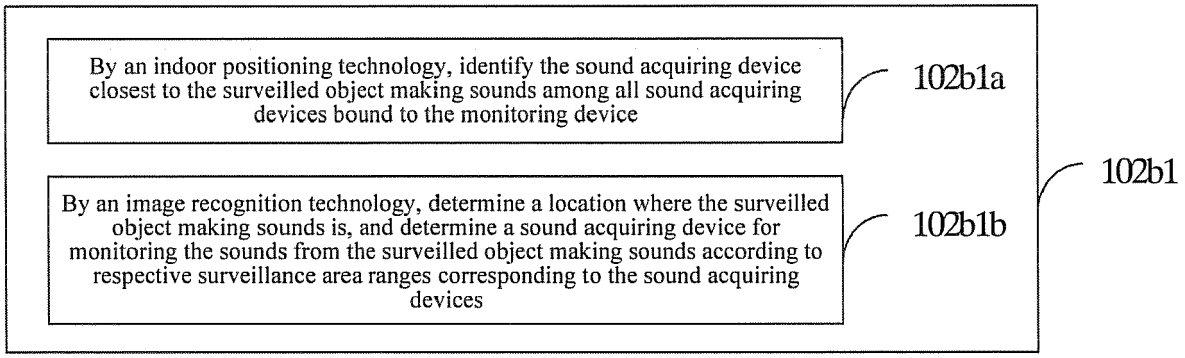

Fig. 2I

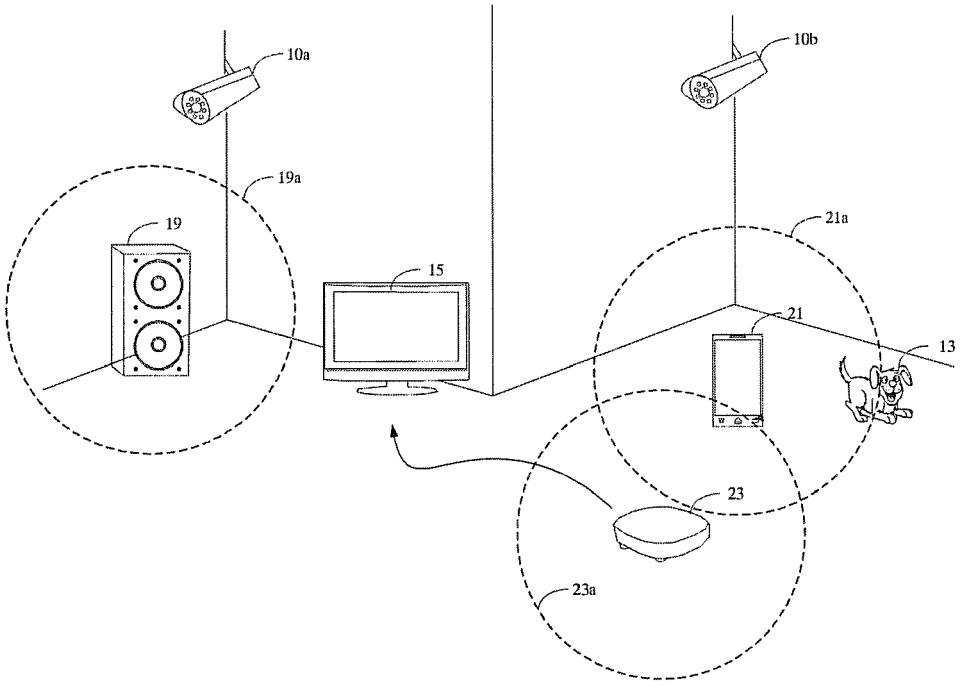
10a
10b
19a
19
15
21a
21
13
23
23a

Fig. 2J

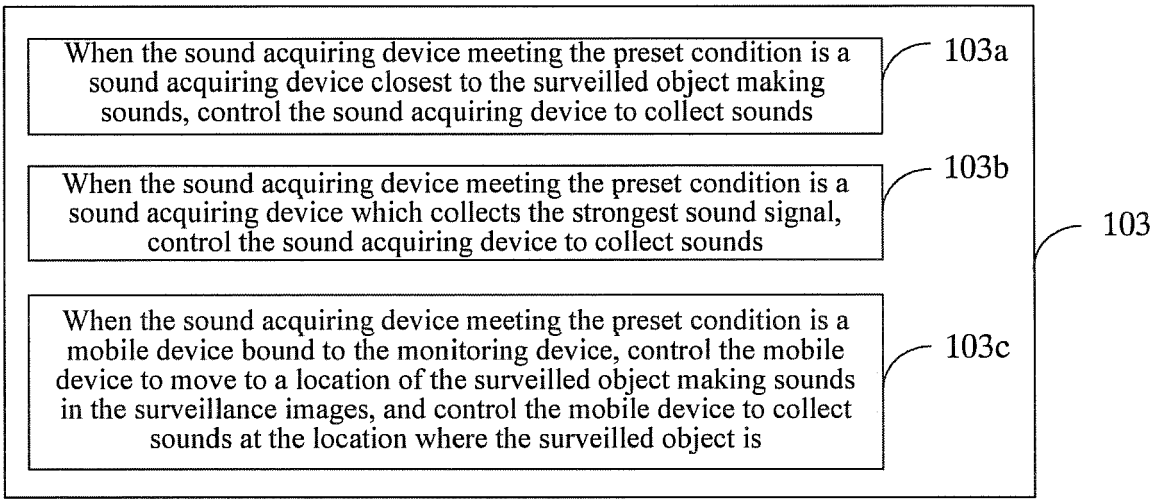
When the sound acquiring device meeting the preset condition is a sound acquiring device closest to the surveilled object making sounds, control the sound acquiring device to collect sounds
103a
When the sound acquiring device meeting the preset condition is a sound acquiring device which collects the strongest sound signal, control the sound acquiring device to collect sounds
103b
When the sound acquiring device meeting the preset condition is a mobile device bound to the monitoring device, control the mobile device to move to a location of the surveilled object making sounds in the surveillance images, and control the mobile device to collect sounds at the location where the surveilled object is
103c
103

Fig. 2K

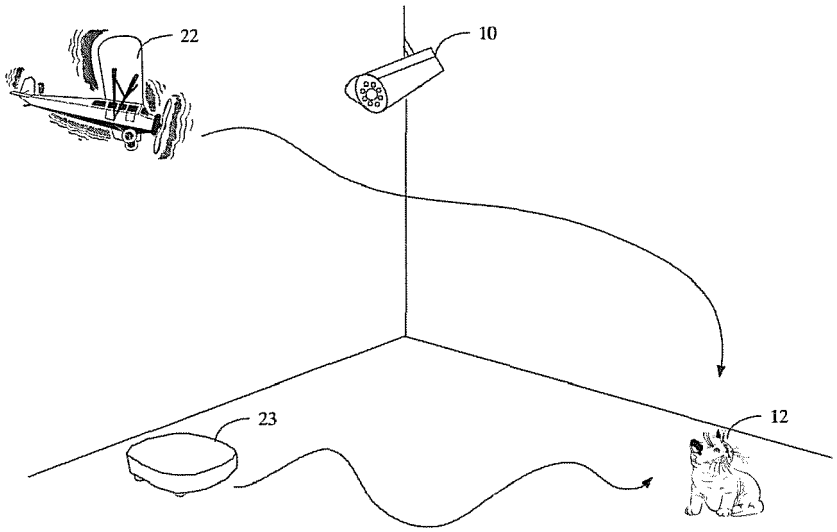
22
10
23
12

Fig. 2L

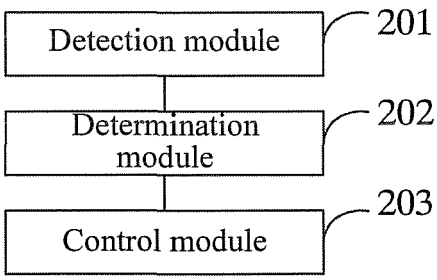
Detection module
201
Determination module
202
Control module
203

Fig. 3

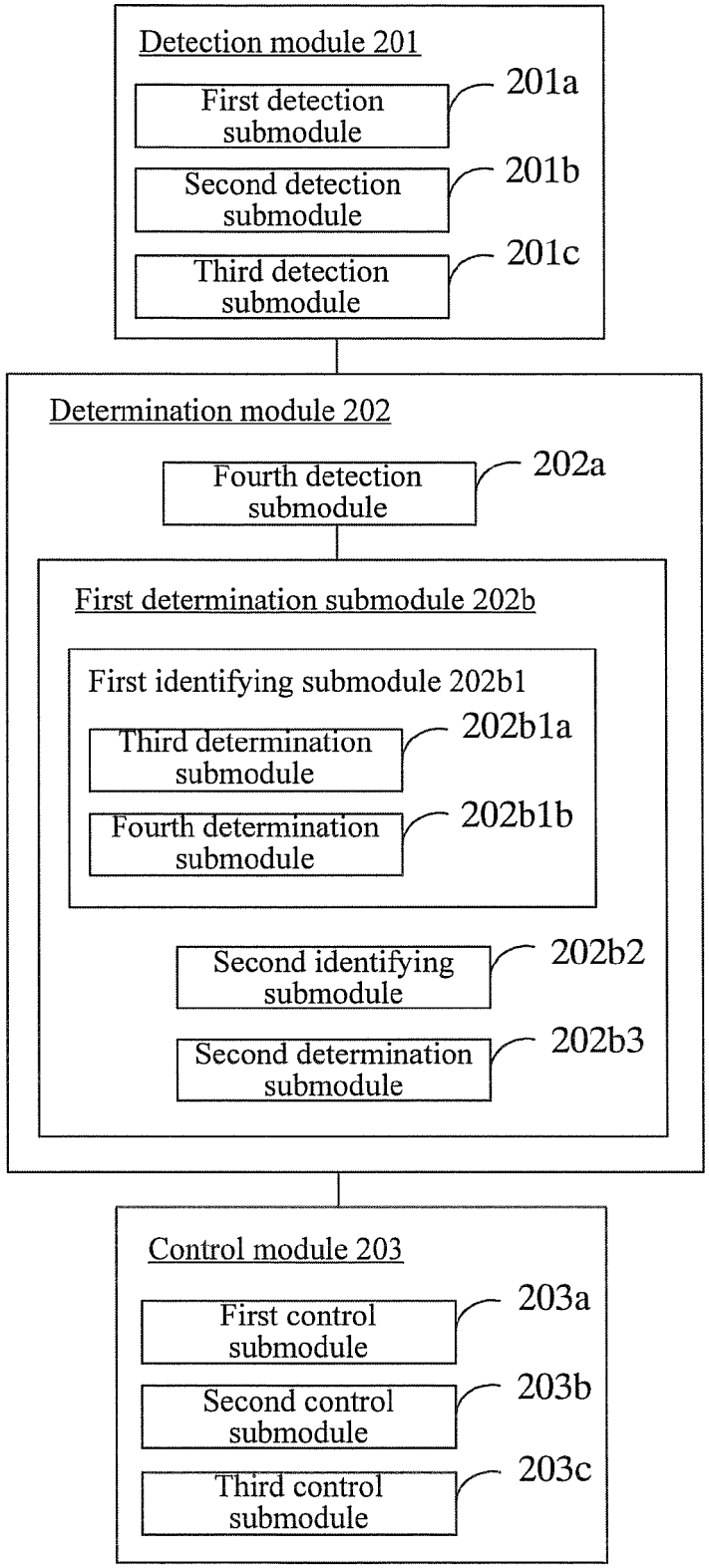
Detection module 201
First detection submodule
201a
Second detection submodule
201b
Third detection submodule
201c
Determination module 202
Fourth detection submodule
202a
First determination submodule 202b
First identifying submodule 202b1
Third determination submodule
202b1a
Fourth determination submodule
202b1b
Second identifying submodule
202b2
Second determination submodule
202b3
Control module 203
First control submodule
203a
Second control submodule
203b
Third control submodule
203c

Fig. 4

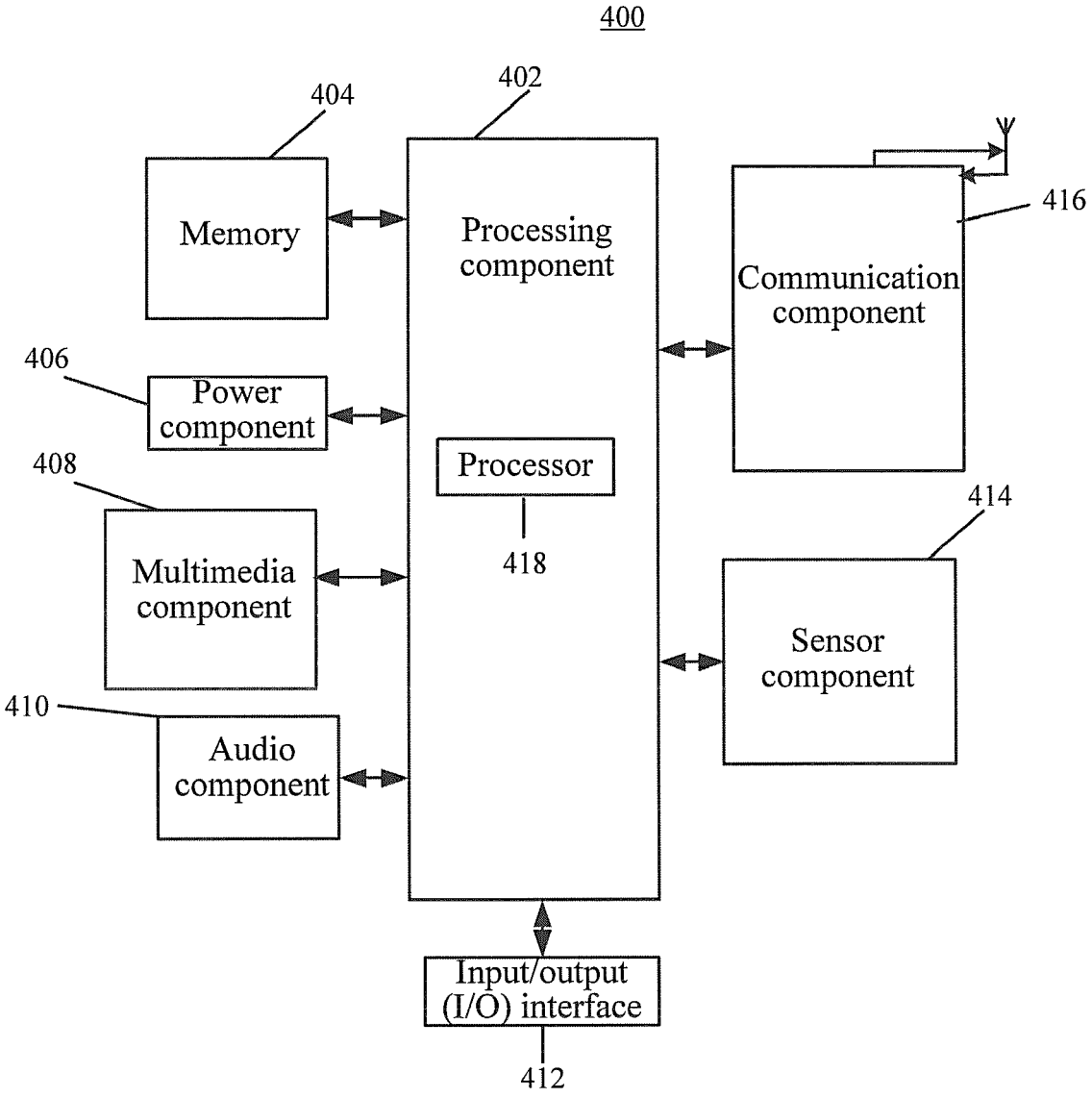
400
402
404
Memory
Processing component
Processor
418
406
Power component
408
Multimedia component
410
Audio component
Communication component
416
414
Sensor component
Input/output (I/O) interface
412

Fig. 5

Europäisches Patentamt
European Patent Office
Office européen des brevets

EUROPEAN SEARCH REPORT

Application Number
EP 16 17 9302

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (IPC)
X	Thomas Jüngling: "Diese Hightech-Polizei schützt Sie vor Einbrechern", WELT.DE, 8 October 2014 (2014-10-08), pages 1-8, XP002765325, Retrieved from the Internet: URL:https://www.welt.de/wirtschaft/webwelt/article133042130/Diese-Hightech-Polizei-schuetzt-Sie-vor-Einbrechern.html [retrieved on 2016-12-15] * the whole document *	1-14	INV. G06K9/00
X	JP 2010 232888 A (IKEGAMI TSUSHINKI KK) 14 October 2010 (2010-10-14) * the whole document *	1-14	
X	KR 2006 0109153 A (S1 CORP [KR]) 19 October 2006 (2006-10-19) * the whole document *	1-14	
A	US 2011/208516 A1 (KUBOYAMA HIDEO [JP]) 25 August 2011 (2011-08-25) * the whole document *	1-14	TECHNICAL FIELDS SEARCHED (IPC) G06K

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
The Hague	16 December 2016	Grigorescu, Cosmin

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

ANNEX TO THE EUROPEAN SEARCH REPORT ON EUROPEAN PATENT APPLICATION NO. EP 16 17 9302

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report. The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

16-12-2016

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
JP 2010232888 A	14-10-2010	JP 5189536 B2	24-04-2013
		JP 2010232888 A	14-10-2010
KR 20060109153 A	19-10-2006	NONE	
US 2011208516 A1	25-08-2011	JP 5489778 B2	14-05-2014
		JP 2011175182 A	08-09-2011
		US 2011208516 A1	25-08-2011

EPO FORM P0459

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82